

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Wahl, Geoffrey M  
O'Gorman, Stephen V

(ii) TITLE OF INVENTION: FLP-MEDIATED GENE MODIFICATION IN  
MAMMALIAN CELLS, AND COMPOSITIONS AND CELLS USEFUL  
THEREFOR

(iii) NUMBER OF SEQUENCES: 4

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark  
(B) STREET: 444 South Flower Street, Suite 2000  
(C) CITY: Los Angeles  
(D) STATE: California  
(E) COUNTRY: USA  
(F) ZIP: 90071

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: US 08/486,409  
(B) FILING DATE: 07-JUN-1995  
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Reiter, Stephen E  
(B) REGISTRATION NUMBER: 31,192  
(C) REFERENCE/DOCKET NUMBER: P41 90004

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (619) 546-1995  
(B) TELEFAX: (619) 546-9392

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1380 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vii) IMMEDIATE SOURCE:

(B) CLONE: NATIVE FLP

(ix) FEATURE:

(A) NAME/KEY: CDS  
(B) LOCATION: 1..1269

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATG	CCA	CAA	TTT	GAT	ATA	TTA	TGT	AAA	ACA	CCA	CCT	AAG	GTG	CTT	GTT	48
Met	Pro	Gln	Phe	Asp	Ile	Leu	Cys	Lys	Thr	Pro	Pro	Lys	Val	Leu	Val	
1				5					10					15		
CGT	CAG	TTT	GTG	GAA	AGG	TTT	GAA	AGA	CCT	TCA	GGT	GAG	AAA	ATA	GCA	96
Arg	Gln	Phe	Val	Glu	Arg	Phe	Glu	Arg	Pro	Ser	Gly	Glu	Lys	Ile	Ala	
			20					25					30			
TTA	TGT	GCT	GCT	GAA	CTA	ACC	TAT	TTA	TGT	TGG	ATG	ATT	ACA	CAT	AAC	144
Leu	Cys	Ala	Ala	Glu	Leu	Thr	Tyr	Leu	Cys	Trp	Met	Ile	Thr	His	Asn	
		35					40					45				
GGA	ACA	GCA	ATC	AAG	AGA	GCC	ACA	TTC	ATG	AGC	TAT	AAT	ACT	ATC	ATA	192
Gly	Thr	Ala	Ile	Lys	Arg	Ala	Thr	Phe	Met	Ser	Tyr	Asn	Thr	Ile	Ile	
	50					55					60					
AGC	AAT	TCG	CTG	AGT	TTC	GAT	ATT	GTC	AAT	AAA	TCA	CTC	CAG	TTT	AAA	240
Ser	Asn	Ser	Leu	Ser	Phe	Asp	Ile	Val	Asn	Lys	Ser	Leu	Gln	Phe	Lys	
65					70					75					80	
TAC	AAG	ACG	CAA	AAA	GCA	ACA	ATT	CTG	GAA	GCC	TCA	TTA	AAG	AAA	TTG	288
Tyr	Lys	Thr	Gln	Lys	Ala	Thr	Ile	Leu	Glu	Ala	Ser	Leu	Lys	Lys	Leu	
				85					90					95		
ATT	CCT	GCT	TGG	GAA	TTT	ACA	ATT	ATT	CCT	TAC	TAT	GGA	CAA	AAA	CAT	336
Ile	Pro	Ala	Trp	Glu	Phe	Thr	Ile	Ile	Pro	Tyr	Tyr	Gly	Gln	Lys	His	
			100					105					110			
CAA	TCT	GAT	ATC	ACT	GAT	ATT	GTA	AGT	AGT	TTG	CAA	TTA	CAG	TTC	GAA	384
Gln	Ser	Asp	Ile	Thr	Asp	Ile	Val	Ser	Ser	Leu	Gln	Leu	Gln	Phe	Glu	
		115					120					125				
TCA	TCG	GAA	GAA	GCA	GAT	AAG	GGA	AAT	AGC	CAC	AGT	AAA	AAA	ATG	CTT	432
Ser	Ser	Glu	Glu	Ala	Asp	Lys	Gly	Asn	Ser	His	Ser	Lys	Lys	Met	Leu	
	130					135					140					
AAA	GCA	CTT	CTA	AGT	GAG	GGT	GAA	AGC	ATC	TGG	GAG	ATC	ACT	GAG	AAA	480
Lys	Ala	Leu	Leu	Ser	Glu	Gly	Glu	Ser	Ile	Trp	Glu	Ile	Thr	Glu	Lys	
145					150					155					160	
ATA	CTA	AAT	TCG	TTT	GAG	TAT	ACT	TCG	AGA	TTT	ACA	AAA	ACA	AAA	ACT	528
Ile	Leu	Asn	Ser	Phe	Glu	Tyr	Thr	Ser	Arg	Phe	Thr	Lys	Thr	Lys	Thr	
				165					170					175		
TTA	TAC	CAA	TTC	CTC	TTC	CTA	GCT	ACT	TTC	ATC	AAT	TGT	GGA	AGA	TTC	576
Leu	Tyr	Gln	Phe	Leu	Phe	Leu	Ala	Thr	Phe	Ile	Asn	Cys	Gly	Arg	Phe	
			180					185					190			
AGC	GAT	ATT	AAG	AAC	GTT	GAT	CCG	AAA	TCA	TTT	AAA	TTA	GTC	CAA	AAT	624
Ser	Asp	Ile	Lys	Asn	Val	Asp	Pro	Lys	Ser	Phe	Lys	Leu	Val	Gln	Asn	
		195					200					205				
AAG	TAT	CTG	GGA	GTA	ATA	ATC	CAG	TGT	TTA	GTG	ACA	GAG	ACA	AAG	ACA	672
Lys	Tyr	Leu	Gly	Val	Ile	Ile	Gln	Cys	Leu	Val	Thr	Glu	Thr	Lys	Thr	
	210					215					220					
AGC	GTT	AGT	AGG	CAC	ATA	TAC	TTC	TTT	AGC	GCA	AGG	GGT	AGG	ATC	GAT	720
Ser	Val	Ser	Arg	His	Ile	Tyr	Phe	Phe	Ser	Ala	Arg	Gly	Arg	Ile	Asp	
225					230					235					240	

CCA CTT GTA TAT TTG GAT GAA TTT TTG AGG AAT TCT GAA CCA GTC CTA	768
Pro Leu Val Tyr Leu Asp Glu Phe Leu Arg Asn Ser Glu Pro Val Leu	
245 250 255	
AAA CGA GTA AAT AGG ACC GGC AAT TCT TCA AGC AAT AAA CAG GAA TAC	816
Lys Arg Val Asn Arg Thr Gly Asn Ser Ser Ser Asn Lys Gln Glu Tyr	
260 265 270	
CAA TTA TTA AAA GAT AAC TTA GTC AGA TCG TAC AAT AAA GCT TTG AAG	864
Gln Leu Leu Lys Asp Asn Leu Val Arg Ser Tyr Asn Lys Ala Leu Lys	
275 280 285	
AAA AAT GCG CCT TAT TCA ATC TTT GCT ATA AAA AAT GGC CCA AAA TCT	912
Lys Asn Ala Pro Tyr Ser Ile Phe Ala Ile Lys Asn Gly Pro Lys Ser	
290 295 300	
CAC ATT GGA AGA CAT TTG ATG ACC TCA TTT CTT TCA ATG AAG GGC CTA	960
His Ile Gly Arg His Leu Met Thr Ser Phe Leu Ser Met Lys Gly Leu	
305 310 315 320	
ACG GAG TTG ACT AAT GTT GTG GGA AAT TGG AGC GAT AAG CGT GCT TCT	1008
Thr Glu Leu Thr Asn Val Val Gly Asn Trp Ser Asp Lys Arg Ala Ser	
325 330 335	
GCC GTG GCC AGG ACA ACG TAT ACT CAT CAG ATA ACA GCA ATA CCT GAT	1056
Ala Val Ala Arg Thr Thr Tyr Thr His Gln Ile Thr Ala Ile Pro Asp	
340 345 350	
CAC TAC TTC GCA CTA GTT TCT CGG TAC TAT GCA TAT GAT CCA ATA TCA	1104
His Tyr Phe Ala Leu Val Ser Arg Tyr Tyr Ala Tyr Asp Pro Ile Ser	
355 360 365	
AAG GAA ATG ATA GCA TTG AAG GAT GAG ACT AAT CCA ATT GAG GAG TGG	1152
Lys Glu Met Ile Ala Leu Lys Asp Glu Thr Asn Pro Ile Glu Glu Trp	
370 375 380	
CAG CAT ATA GAA CAG CTA AAG GGT AGT GCT GAA GGA AGC ATA CGA TAC	1200
Gln His Ile Glu Gln Lys Gly Ser Ala Glu Gly Ser Ile Arg Tyr	
385 390 395 400	
CCC GCA TGG ATT GGG ATA ATA TCA CAG GAG GTA CTA GAC TAC CTT TCA	1248
Pro Ala Trp Ile Gly Ile Ile Ser Gln Glu Val Leu Asp Tyr Leu Ser	
405 410 415	
TCC TAC ATA AAT AGA CGC ATA TAAGTACGCA TTTAAGCATA AACACGCACT	1299
Ser Tyr Ile Asn Arg Arg Ile	
420	
ATCCCGTTCT TCTCATGTAT ATATATATAC AGGCAACACG CAGATATAGG TGCGACGTGA	1359
ACAGTGAGCT GTATGTGCGC A	1380

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 423 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

2.2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Pro	Gln	Phe	Asp	Ile	Leu	Cys	Lys	Thr	Pro	Pro	Lys	Val	Leu	Val	1	5	10	15
Arg	Gln	Phe	Val	Glu	Arg	Phe	Glu	Arg	Pro	Ser	Gly	Glu	Lys	Ile	Ala	20	25	30	
Leu	Cys	Ala	Ala	Glu	Leu	Thr	Tyr	Leu	Cys	Trp	Met	Ile	Thr	His	Asn	35	40	45	
Gly	Thr	Ala	Ile	Lys	Arg	Ala	Thr	Phe	Met	Ser	Tyr	Asn	Thr	Ile	Ile	50	55	60	
Ser	Asn	Ser	Leu	Ser	Phe	Asp	Ile	Val	Asn	Lys	Ser	Leu	Gln	Phe	Lys	65	70	75	80
Tyr	Lys	Thr	Gln	Lys	Ala	Thr	Ile	Leu	Glu	Ala	Ser	Leu	Lys	Lys	Leu	85	90	95	
Ile	Pro	Ala	Trp	Glu	Phe	Thr	Ile	Ile	Pro	Tyr	Tyr	Gly	Gln	Lys	His	100	105	110	
Gln	Ser	Asp	Ile	Thr	Asp	Ile	Val	Ser	Ser	Leu	Gln	Leu	Gln	Phe	Glu	115	120	125	
Ser	Ser	Glu	Glu	Ala	Asp	Lys	Gly	Asn	Ser	His	Ser	Lys	Lys	Met	Leu	130	135	140	
Lys	Ala	Leu	Leu	Ser	Glu	Gly	Glu	Ser	Ile	Trp	Glu	Ile	Thr	Glu	Lys	145	150	155	160
Ile	Leu	Asn	Ser	Phe	Glu	Tyr	Thr	Ser	Arg	Phe	Thr	Lys	Thr	Lys	Thr	165	170	175	
Leu	Tyr	Gln	Phe	Leu	Phe	Leu	Ala	Thr	Phe	Ile	Asn	Cys	Gly	Arg	Phe	180	185	190	
Ser	Asp	Ile	Lys	Asn	Val	Asp	Pro	Lys	Ser	Phe	Lys	Leu	Val	Gln	Asn	195	200	205	
Lys	Tyr	Leu	Gly	Val	Ile	Ile	Gln	Cys	Leu	Val	Thr	Glu	Thr	Lys	Thr	210	215	220	
Ser	Val	Ser	Arg	His	Ile	Tyr	Phe	Phe	Ser	Ala	Arg	Gly	Arg	Ile	Asp	225	230	235	240
Pro	Leu	Val	Tyr	Leu	Asp	Glu	Phe	Leu	Arg	Asn	Ser	Glu	Pro	Val	Leu	245	250	255	
Lys	Arg	Val	Asn	Arg	Thr	Gly	Asn	Ser	Ser	Ser	Asn	Lys	Gln	Glu	Tyr	260	265	270	
Gln	Leu	Leu	Lys	Asp	Asn	Leu	Val	Arg	Ser	Tyr	Asn	Lys	Ala	Leu	Lys	275	280	285	
Lys	Asn	Ala	Pro	Tyr	Ser	Ile	Phe	Ala	Ile	Lys	Asn	Gly	Pro	Lys	Ser	290	295	300	
His	Ile	Gly	Arg	His	Leu	Met	Thr	Ser	Phe	Leu	Ser	Met	Lys	Gly	Leu	305	310	315	320

23

Thr	Glu	Leu	Thr	Asn	Val	Val	Gly	Asn	Trp	Ser	Asp	Lys	Arg	Ala	Ser
				325					330					335	
Ala	Val	Ala	Arg	Thr	Thr	Tyr	Thr	His	Gln	Ile	Thr	Ala	Ile	Pro	Asp
			340					345					350		
His	Tyr	Phe	Ala	Leu	Val	Ser	Arg	Tyr	Tyr	Ala	Tyr	Asp	Pro	Ile	Ser
		355					360					365			
Lys	Glu	Met	Ile	Ala	Leu	Lys	Asp	Glu	Thr	Asn	Pro	Ile	Glu	Glu	Trp
	370					375					380				
Gln	His	Ile	Glu	Gln	Leu	Lys	Gly	Ser	Ala	Glu	Gly	Ser	Ile	Arg	Tyr
385					390					395					400
Pro	Ala	Trp	Ile	Gly	Ile	Ile	Ser	Gln	Glu	Val	Leu	Asp	Tyr	Leu	Ser
				405					410					415	
Ser	Tyr	Ile	Asn	Arg	Arg	Ile									
			420												

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 34 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GAAGTTCCTA TTCTCTAGAA AGTATAGGAA CTTC

34

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 68 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GATCCCGGGC TACCATGGAG AAGTTCCTAT TCCGAAGTTC CTATTCTCTA GAAAGTATAG  
GAACTTCA

60

68

34